Realising traceable electrostatic forces despite non-linear balance motion

Julian Stirling\textsuperscript{1,2} and Gordon A. Shaw\textsuperscript{1}
\textsuperscript{1} National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20851, USA
\textsuperscript{2} Joint Quantum Institute, University of Maryland, College Park, MD 20742, USA
E-mail: julian.stirling@nist.gov

Abstract. Direct realisation of force, traceable to fundamental constants via electromagnetic balances, is a key goal of the proposed redefinition of the International System of Units (SI). This will allow small force metrology to be performed using an electrostatic force balance (EFB) rather than subdivision of larger forces. Such a balance uses the electrostatic force across a capacitor to balance an external force. In this paper we model the capacitance of a concentric cylinder EFB design as a function of the displacement of its free electrode, accounting for the arcuate motion produced by parallelogram linkages commonly used in EFB mechanisms. From this model we suggest new fitting procedures to reduce uncertainties arising from non-linear motion as well as methods to identify misalignment of the mechanism. Experimental studies on both a test capacitor and the NIST EFB validate the model.

Keywords: Metrology, Electrostatics, Precision measurement, Mass, Force


1. Introduction

Under the proposed redefinition of the International System of Units (SI), the kilogram will be defined in terms of three fundamental constants (the Plank constant, the speed of light, and the frequency of the radiation corresponding to the transition between the two hyperfine levels of the Caesium-133 atom in its ground state). This will allow measurement of SI traceable mass or force in the laboratory without the use of transfer standards used to link to the International Kilogram Prototype.

This link to fundamental constants has the potential to revolutionise small mass and force metrology. Under the current system, multiple subdivisions from the kilogram must be made to produce small mass transfer standards. Each of these subdivisions adds further compounding uncertainties, progressively limiting the attainable accuracy of mass and force measurements as they decrease in magnitude, with potential adverse consequences
Realising traceable electrostatic forces despite non-linear balance motion for nanometer-scale science and technology. Once the kilogram is redefined, SI traceable force can be directly realised in the laboratory at any scale, measurements will again be limited by the accuracy of the experiment rather than our knowledge of the unit itself.

The standard method for realising mass directly is the Watt balance. This uses the Lorentz force on a coil of wire in a magnetic field to balance the gravitational force of a test mass by adjusting the current through the wire. Electrostatic forces can also be used to realise force directly from the proposed redefined SI. In practice, large electrostatic forces are more difficult to generate than electromagnetic forces. Thus, at the scale of the kilogram an electrostatic force balance (EFB) is not practical. At small forces a higher relative uncertainty is acceptable, and thus equipment requires less complicated metrology. As such, accurate electrostatic force balances have the potential to realise SI traceable masses and forces at previously unattainable scales. This principle has been used for both small mass metrology [1] and for measuring Newton’s gravitational constant [2].

The basic principle of an EFB is to balance the force under study with a known electrostatic force. The electrostatic force is generated between two electrodes of a capacitor of capacitance $C$. The force depends on the gradient of the capacitance $F = \frac{1}{2}V^2\nabla C$, \hspace{1cm} (1)

where $V$ is the voltage across the capacitor. A balance mechanism can be designed so the motion of the two plates is constrained to one dimension, $z$. Then Equation 1 simplifies to

$$F_z = \frac{1}{2}V^2 \frac{\partial C}{\partial z}. \hspace{1cm} (2)$$

By choosing an appropriate capacitor geometry the capacitance gradient in the $z$ direction can be linear. For example a concentric cylinder capacitor where the cylinder axes are aligned with the $z$-axis.

The sources of uncertainty in such a balance are the accuracy with which the voltage and the capacitance gradient can be measured, along with any further uncertainties introduced by the balance mechanism and its kinematics. To realise mass rather than force the local gravity must also be measured. This can be done, however, using an absolute gravimeter with relative uncertainties of order $10^{-8}$.

Attainable uncertainties in voltage, capacitance, and distance will not be the limiting factors in small force metrology. Voltage realisations can achieve relative uncertainties of order $10^{-11}$ [3], by direct link to quantum effects. Using the calculable capacitor, capacitance can be realised with relative uncertainties of $15 \times 10^{-9}$ [4]. Interferometric distance measurement achieve absolute uncertainties measured in attometers [5] for atomic scale measurements, and relative uncertainties $4 \times 10^{-10}$ for a range of a few centimeters [6]. In practice, a balance designed to operate at the 1 mg ($\approx 10 \mu$N) level—approximately 6 orders of magnitude lower force than the Watt balance—requires a capacitor where $\frac{\partial C}{\partial z}$ is 1 pF/mm to operate at 100 V. At these scales commercially available equipment [7]—used
in the National Institute of Standards and Technology (NIST) EFB—can achieve relative uncertainties of $5 \times 10^{-6}$, $15 \times 10^{-8}$, and $4 \times 10^{-6}$ for capacitance [8], distance [9], and voltage, respectively [10]. Thus, if further uncertainties could be avoided this would allow realisation of the milligram with an uncertainty of order $1 \times 10^{-5}$. By further calibration of the multimeter and capacitance bridge, combined with careful experimental design these uncertainties could be reduced an order of magnitude further.

Current calibrations for the milligram, relying on subdivision from a kilogram artifact, have relative uncertainties of order $2 \times 10^{-4}$ [11]. Thus, an electrostatic force balance could reduce uncertainties at this level by two orders of magnitude, provided no additional uncertainties accrue. Both Type A uncertainties from sources such as the stability of the control loop used to control the balance position, and Type B uncertainties from issues such as alignment need to be considered in such a measurement. In this paper we concentrate on Type B uncertainties resulting from the alignment of the electrodes and the electrode translation mechanism, and how this relates to the ability to measure the capacitance gradient. For this we analyse a concentric cylinder capacitor such as used in the NIST electrostatic force balance [1].

2. Theory

As previously mentioned, a cylindrical capacitor design can be used to generate electrostatic forces. To remove any off-axis forces the electrodes should be concentric. For a truly concentric cylinder capacitor with electrodes overlapping by $H$ in the $z$ (axial) direction, the capacitance per unit length can be calculated as

$$C = \frac{\epsilon_r}{2k_e \log \left( \frac{a_2}{a_1} \right)}.$$  \hfill (3)

where $a_1$ and $a_2$ are the radii of the inner and outer cylinder respectively. $k_e$ is the Coulomb constant ($= \frac{1}{4\pi \epsilon_0}$), and $\epsilon_r$ is the relative permittivity of the medium between the electrodes.

For an accurate force measurement two different readings must be taken. Firstly, the capacitance gradient should be measured by recording the capacitance as the balance position is varied. Secondly a differential recording of the voltage measurement is performed, to measure the difference in the voltage needed to maintain the balance’s null point with and without an applied force. These two results can be combined to a final force measurement using Equation 2.

This method assumes the free electrode remains axially aligned and concentric with the fixed electrode throughout its entire range of motion, resulting in a linear capacitance gradient. If this is not the case higher-order terms will be present in $C(z)$, the capacitance with respect to position. If force measurements are performed differentially at $z = 0$ then they are insensitive to any higher order terms. The capacitance gradient of the balance at its null position ($C'(0)$), however, is determined from capacitance measurements at a range
Realising traceable electrostatic forces despite non-linear balance motion

Figure 1. a) Schematic of the NIST electrostatic force balance. The capacitor used to generate traceable electrostatic forces is attached to a parallelogram linkage. The schematic also shows the tension spring used to control the balance stiffness and the auxiliary capacitor used to control the balance motion when measuring the capacitance gradient of the main capacitor. b) Diagram showing the arcuate motion of the inner cylinder in ideal alignment. \( L \) is the length of the linkage arms, \( \alpha \) is the angle of the linkage, and \( H \) is the \( z \)-direction overlap of the electrodes when \( z = 0 \). c) Diagram showing possible misalignment of the capacitor cylinders. \( b_0 \) and \( \alpha_0 \) are the distance between the electrode axes and the angle of the balance arm at \( z = 0 \) respectively.

of \( z \) positions. Any higher terms in the capacitance vs. \( z \)-position must be accounted for when fitting the data to correctly determine the capacitance gradient at \( z = 0 \). As such it is essential to know the functional form of \( C(z) \).

In practice, a balance mechanism which constrains the electrodes to be both axially aligned and concentric throughout its motion is a considerable challenge. A standard pan balance design uses a balance beam with the balance pans freely-suspended; gravity maintains the balance pan’s angle as the balance beam tilts. For an EFB free suspension is not possible; while the off axial forces on a truly concentric capacitor are zero, a freely suspended electrode would be in an unstable equilibrium. This problem is solved by rigidly attaching the electrode to a parallelogram linkage. The parallelogram linkage keeps the axial alignment of the cylinders constant throughout its motion. Both this linkage and a balance beam cannot keep the electrodes concentric, instead there would be lateral displacement proportional to the cosine of the balance tilt. For other types precision balances, such as
a Watt balance, lateral displacement is avoided by suspending balance pans from a band attached to a balance wheel. For an EFB such a mechanism would again introduce an unstable equilibrium.

Both the NIST EFB [1] and other EFBs [12] implement the parallelogram linkage as the cylinder can be aligned to be concentric at the null position maintained during force measurements. In this paper we consider the functional form of $C(z)$ arising from the resulting arcuate motion of the inner electrode in an aligned state, and the effects of misalignment (Figures 1a and b respectively). An aligned state is defined at $z = 0$ when the balance arms are horizontal, and the electrodes are concentric.

Considering first the effect of eccentricity. If the central axes of the two electrodes are separated by a distance $b$ the capacitance per unit length becomes [13]

$$\frac{C}{H} = \frac{\epsilon_r}{2k_e} \log \left( \frac{\sqrt{((a_2-a_1)^2-b^2)((a_2+a_1)^2-b^2)} - b^2 + a_2^2 + a_1^2}{2a_1a_2} \right),$$

(4)

This form can be shown to be consistent with that given in Smythe [14] by considering that $\text{arcosh}(x) = \log(x + \sqrt{x^2 - 1})$, and in the case of $b = 0$ reduces to Equation 3.

Taking the Taylor expansion of Equation 4 about $b = 0$ gives

$$\frac{C}{H} = \frac{\epsilon_r}{2k_e} \left( \frac{1}{\log \left( \frac{a_2}{a_1} \right)} + \frac{b^2}{\xi} + O(b^4) \right),$$

(5)

where $\xi = (a_2 - a_1)(a_2 + a_1) \log^2 \left( \frac{a_2}{a_1} \right)$. This series converges quickly for $a_1 \gtrsim (a_2 - a_1)$ and $b^2 \ll \frac{4(a_2-a_1)^2}{3}$.

Defining $H$ as the z-direction overlap at $z = 0$, in the aligned case, Equation 4 simply expands to

$$C(z) = C_0 + \frac{\epsilon_r}{2k_e} \left( \frac{z}{\log \left( \frac{a_2}{a_1} \right)} - \frac{Hz^4 + z^5}{4L^2 \xi} + O(z^8) \right),$$

(6)

by also using the first two terms in the expansion for $b$ in terms of $z$. Here $C_0$ is the capacitance at $z = 0$. We note that despite the fifth order functional form no second or third order terms are present. However, introducing the misalignments $b_0$ and $\alpha_0$ from Figure 1b results in a significantly more convoluted functional form:

$$C(z) = C_0 + \frac{\epsilon_r}{2k_e} \left( \frac{z}{\log \left( \frac{a_2}{a_1} \right)} - \frac{(b_0^2 + 2HB_0 \tan(\alpha_0))z + (2b_0 \tan(\alpha_0) + H \tan^2(\alpha_0))z^2 + \tan^2(\alpha_0)z^3}{\xi} \right.$$

$$+ \frac{Hb_0z^2 - (H \tan(\alpha_0) - b_0)z^3 - \tan(\alpha_0)z^4}{L \cos(\alpha_0) \xi} - \frac{Hz^4 + z^5}{4L^2 \cos^2(\alpha_0) \xi} + O(z^6) \bigg).$$

(7)

Both $C_0$ and the first order term depend on $b_0$ and $\alpha_0$, and these misalignments also introduce second and third order terms.
3. Experimental

The appropriate order of a polynomial fit to a capacitance gradient could be determined empirically by observing the residuals arising from lower order fits. It is, however, disconcerting to apply high order polynomial fits to data without good theoretical grounding for the higher terms. With enough free parameters one can fit a curve to any data set. One could also use orthogonal polynomials to produce a more robust fit at lower order. The disadvantage of this for the EFB is the gradient of the higher order terms is no-longer zero at \( z = 0 \). As such all terms must be combined to determine the gradient at \( z = 0 \), rather than the first order term and its associated uncertainty from covariance.

To experimentally validate the necessity for higher order fits when determining \( C'(0) \) we have constructed a capacitor from a pair of cylindrical electrodes. The inner electrode was held at ground and the outer electrode was fitted with a grounding guard electrode on top to limit stray capacitance. The diameters of the inner and outer electrodes were nominally 22 mm, and 23 mm respectively. The cylindrical electrodes were manufactured with a reflective surface parallel to the central axis to a machine tolerance of 10 \( \mu \)rad. The inner electrode was attached to a Newport 562-XYZ 3-axis stage, controlled with three Newport VP-25AA high-precision motorised actuators with 100 nm resolution, to simulate balance motion. An Andeen-Hagerling 2700A bridge measured capacitance between the electrodes at a frequency of 1 kHz and a maximum potential of 15 V. All measurements were performed at a temperature of 20 ± 0.5 °C, a pressure of 1 atmosphere, and a relative humidity of 40%.

The angle between the electrodes was measured using a simple optical lever technique (Figure 2). A divergent laser beam was split between mirrors on the inner and outer electrode, the reflected beams were projected on to a screen 2.69 m from the mirror of the outer electrode. At the screen the laser spot sizes were \( \approx 7 \) mm in diameter. As the distance from the inner electrode mirror to the outer electrode mirror \( (q) \) was varied between 80 mm...
Table 1. Experimental parameters with Type A uncertainties ($k = 1$)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_0$</td>
<td>19.846161(61) pF</td>
<td>Calculated fit</td>
</tr>
<tr>
<td>$C'$</td>
<td>1.116902(32) nF/m</td>
<td></td>
</tr>
<tr>
<td>$a_1$</td>
<td>11.023(4) mm</td>
<td>Measured with encoded translation stage.</td>
</tr>
<tr>
<td>$a_1/a_2$</td>
<td>1.051105(30)</td>
<td>Using Eq. 3.</td>
</tr>
<tr>
<td>$a_2$</td>
<td>11.586(4) mm</td>
<td>From $a_1$ and $a_1/a_2$.</td>
</tr>
</tbody>
</table>

and 64 mm, no visible changes to the spot positions were recorded. The angle between the incident and reflected beams was $3^\circ$, thus this should separate the final spots by $\approx 4$ mm. For the optical lever length used, a systematic spot separation of 4 mm results in a 1.5 mrad misalignment. Taking into account spot size, an uncertainty of 3 mrad was associated with angular alignment.

Before each sweep in $z$, the inner electrode was centred in $x$ and $y$ at $z = -8$ mm and

![Figure 3](image-url)

**Figure 3.** a) Capacitance gradient for linear motion of the inner electrode. These results were used to verify the alignment of the test setup and to determine the capacitance gradient. b) Residuals from a first order fit are presented. These show some minimal sign of a parabolic structure which indicate slight misalignment of the cylinder axes. No residual are greater than 500 aF. For comparison a 3 mrad misalignment of the motion axis would result in a 520 aF deviation from linearity at $z = 3.2$ mm. Error bars represent one standard deviation from repeated measurements.
8 mm by minimising the capacitance, as described previously. The central z axis was then defined between these two points; from this axis the motion to sweep in x, y, and z was calculated. As large motions in z amplify unintended misalignments, higher order terms, and capacitor edge effects, sweep data from \( z = \pm 3.2 \) mm was analysed. This 6.4 mm range of motion is still large compared to the range of motion of the NIST EFB which is approximately 2.2 mm.

As an initial test of the experimental set-up and alignment, a capacitance gradient was taken using linear motion with the central axes of the two electrodes aligned. The data, shown in Figure 3, was fitted with a first order polynomial, \( C = C_0 + C'z \). The fit’s adjusted coefficient of determination differs from unity by a few parts per billion \( (1 - R^2 = 1.4 \times 10^{-8}) \). All data points lie within 500 attofarad of the fit. The fit results and derived parameters are shown in Table 1.

First considering the aligned case where at \( z = 0 \) the electrodes are concentric and the arc angle is zero (Figure 1a); no second or third order polynomial terms should be present in \( C(z) \). Without changing the alignment from the linear measurements the inner electrode was set to move over an arc with a maximum deviation of 0.2 mm at \( \pm 8 \) mm. This is consistent

![Figure 4](image_url)

**Figure 4.** a) Fit residuals for a \( C(z) \) measured with arcuate balance motion. Fits performed for different polynomial models. b) First order capacitance gradient, \( C'(0) \), extracted from each fit. Error bars represent \( k = 1 \) type A uncertainty estimated from least squares fitting. The horizontal line and grey area represent the capacitance gradient and type A uncertainty from the linear motion presented in Figure 3.
with a beam length of $L = 160.1$ mm for a 4-bar linkage.

One method to validate the model would be to use the parameters in Table 1 to check the magnitude of higher order terms. However, the fit parameters for the higher order terms are very sensitive to any measurement uncertainties, and the accuracy to which they can be measured is of little interest for electrostatic forces generated at $z = 0$. Thus, we instead fit our model and the first five orders of polynomial fits to the data and compare the returned value for the gradient at $z = 0$ to the gradient measured for the linear motion (Figure 4).

From Figure 4 it is clear that the model of fourth and fifth order terms added to the linear capacitance gradient (henceforth, referred to as a reduced-fifth-order polynomial) returns the same result as linear motion. The odd-order polynomial fits also return the correct result within the uncertainty from the linear least squares fit, but with significantly larger uncertainty than the reduced-fifth-order polynomial model. The second and fourth order polynomial fits, however, return values more than two sigma from the linear result. It is worth noting that the residuals of the fourth order polynomial are not significantly larger than that of the reduced-fifth-order polynomial nor the fifth order polynomial, nor do they display a clear structure.

Our model shows that the introduction of an offset in the $x-y$ plane, such that the two electrodes are not concentric at $z = 0$, introduces second and third order terms in $z$ to the capacitance sweep. Also the capacitance gradient at $z = 0$ will increase as the electrodes are no longer concentric, the condition for a capacitance gradient minimum. For this test a smaller amplitude of the arcuate motion was used, in this case corresponding to a parallelogram linkage with length $L = 320.05$ mm. The centre of the motion was laterally offset by $\pm 50 \mu$m. As expected, Figure 5 shows that the reduced-fifth-order polynomial model is no longer a good fit to the data once the offset is introduced. As such a full fifth-order polynomial is necessary to account for possible misalignments. Inspecting the reduced-fifth-order fit residuals could potentially lead to a method for checking misalignment. However, misalignments from other sources, such as $\alpha_0 \neq 0$ will also cause similar residuals, thus determining the exact source of misalignment may be difficult in practice.

4. Discussion

The presented results agree with the theoretical model, and confirm that the arcuate motion induced by parallelogram linkages on EFBs can be accurately accounted for if the correct model is used to when fitting $C(z)$. Comparing uncertainties in $C'(0)$ for our test setup for linear and arcuate motion shows that the relative uncertainty increases from $29 \times 10^{-6}$ in the linear case to $66 \times 10^{-6}$ and $143 \times 10^{-6}$ for the reduced-fifth-order polynomial and fifth-order polynomial fits respectively for arcuate motion. The increased uncertainty arises from the extra degrees of freedom of the fit.

It is important to note that the magnitude of the relative $C'(0)$ uncertainties presented in this paper are not of particular significance. The experiment was designed to easily compare linear and non-linear motion, rather than to produce the lowest uncertainty. As such it
is worth considering results from the NIST EFB. In Figure 6a we compare the calculated values for \( C'(0) \) from 6 capacitance gradient determinations taken over 12 days in vacuum on the EFB. Capacitance was measured at 11 discrete \( z \) positions over a nominal range of \( \pm 1 \) mm. To avoid polarisation mixing errors in the double-pass Michelson interferometer used to measure displacement, each \( z \) position used is a multiple of one quarter of the wavelength of the laser. Capacitance readings were performed with an Andeen-Hagerling 2500A bridge traceably calibrated to the NIST calculable capacitor. For more details on the instrumentation of the NIST EFB see \citep{1, 15}.

The data in Figure 6a was not fitted with the reduced-fifth-order polynomial as the residuals plotted in Figure 6b make it clear that the second and third order terms are needed to correct for misalignment. These residuals are qualitatively very similar to those presented in Figure 5 except notably the amplitude is almost three orders of magnitude smaller, with a maximum absolute residual of only 21 aF in for the reduced-fifth-order case, and 2 aF for the fifth order fit. This approaches the 0.5 aF resolution of the capacitance bridge.

Figure 6a compares odd order polynomial fitting up to the suggested fifth order fit.
Realising traceable electrostatic forces despite non-linear balance motion

The results reveal a relative systematic decrease in $C''(0)$ of approximately $30 \times 10^{-6}$ when using fifth order polynomial fitting rather than a simple linear fit. This discrepancy exceeds type A uncertainty from the linear fit at $5\sigma$. Third order fitting compared to 5th results in a relative effect of $8 \times 10^{-6}$ which, while outside the type A uncertainty for the measurement, is smaller than the measurement drift between results. While a relative uncertainty of $30 \times 10^{-6}$ may be dominant for some EFB results, this effect is highly dependent on the distance the electrodes are swept, and the electrode alignment, so should not be considered a systematic uncertainty which can retroactively be applied.

5. Conclusion

We have derived a model for concentric cylinder electrostatic force balances using parallelogram linkages and verified it experimentally. The model shows 5th order polynomial terms in the capacitance gradient used to realise a traceable force, and that lower order fits can result in a systematic error in the measured force. The magnitude of this error
Realising traceable electrostatic forces despite non-linear balance motion

depends on a number of factors such as the dimensions of the instrument, the distance moved in the capacitance gradient determination, and the balance alignment. For measurements performed on the NIST EFB the higher order fits resulted in a relative reduction of \( C'(0) \) (and hence the measured force) by \( 30 \times 10^{-6} \) compared to a linear fit. We note that this should not be considered a fixed systematic for the apparatus as it depends on both the current alignment and the displacement over which the capacitance is measured. It does, however, confirm that as future work pushes micro/nano Newton force metrology to the parts per million level the full functional form of the \( C(z) \) must be considered.

6. References


[7] Certain commercial equipment, instruments, or materials (or suppliers, or software, etc.) are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.


